Exceptional rotations of random graphs: a VC theory

Louigi Addario-Berry [*]   Shankar Bhamidi [†]   Sébastien Bubeck [‡]   Luc Devroye [§]

Gábor Lugosi [¶]   Roberto Imbuzeiro Oliveira [‖]

June 10, 2015

Abstract

In this paper we explore maximal deviations of large random structures from their typical behavior. We introduce a model for a high-dimensional random graph process and ask analogous questions to those of Vapnik and Chervonenkis for deviations of averages: how "rich" does the process have to be so that one sees atypical behavior.

In particular, we study a natural process of Erdős-Rényi random graphs indexed by unit vectors in $\mathbb{R}^d$. We investigate the deviations of the process with respect to three fundamental properties: clique number, chromatic number, and connectivity. In all cases we establish upper and lower bounds for the minimal dimension d that guarantees the existence of "exceptional directions" in which the random graph behaves atypically with respect to the property. For each of the three properties, four theorems are established, to describe upper and lower bounds for the threshold dimension in the subcritical and supercritical regimes.

1   Introduction

One of the principal problems in probability and statistics is the understanding of maximal deviations of averages from their means. The revolutionary work of Vapnik and Chervonenkis [19, 20, 21] introduced a completely new combinatorial approach that opened many paths and helped us understand this fundamental phenomena. Today, the Vapnik-Chervonenkis theory has become the theoretical basis of statistical machine learning, empirical process theory, and has applications in a diverse array of fields.

The purpose of this paper is to initiate the exploration of maximal deviations of complex random structures from their typical behavior. We introduce a model for a high-dimensional random graph process and ask analogous questions to those of Vapnik and Chervonenkis for deviations of averages: how "rich" does the process have to be so that one sees atypical behavior. In particular, we study a process of Erdős-Rényi random graphs. In the $G(n,p)$ model introduced by Erdős-Rényi [8, 7], a graph on n vertices is obtained by connecting each pair of vertices with probability p, independently,

---

[*]Department of Mathematics and Statistics, McGill University; louigi.addario@mcgill.ca

[†]Department of Statistics and Operations Research, University Of North Carolina, Chapel Hill; bhamidi@email.unc.edu

[‡]Microsoft Research and Princeton University; sebubeck@microsoft.com

[§]School of Computer Science, McGill University; lucdevroye@gmail.com

[¶]ICREA and Pompeu Fabra University, Barcelona; gabor.lugosi@upf.edu

[‖]IMPA, Rio de Janeiro, Brazil; rimfo@impa.br

at random. The $G(n,p)$ model has been thoroughly studied and many of its properties are well understood—see, e.g., the monographs of Bollobás [4] and Janson, Łuczak, and Ruciński [9].

In this paper we introduce a random graph process indexed by unit vectors in $\mathbb{R}^d$, defined as follows. For positive integer n, write $[n] = \{1, \ldots, n\}$. For $1 \leq i < j \leq n$, let $X_{i,j}$ be independent standard normal vectors in $\mathbb{R}^d$. Denote by $\boldsymbol{X}_n = (X_{i,j})_{1 \leq i < j \leq n}$ the collection of these random points. For each $s \in S^{d-1}$ (where S^{d-1} denotes the unit sphere in $\mathbb{R}^d$) and $t \in \mathbb{R}$ we define the random graph $\Gamma(\boldsymbol{X}_n, s, t)$ with vertex set $v(\Gamma(\boldsymbol{X}_n, s, t)) = [n]$ and edge set $e(\Gamma(\boldsymbol{X}_n, s, t)) = \{\{i,j\} : \langle X_{i,j}, s \rangle \geq t\}$, where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathbb{R}^d$.

For any fixed $s \in S^{d-1}$ and $t \in \mathbb{R}$, $\Gamma(\boldsymbol{X}_n, s, t)$ is distributed as an Erdős-Rényi random graph $G(n,p)$, with $p = 1 - \Phi(t)$ where Φ is the distribution function of a standard normal random variable. In particular, $\Gamma(\boldsymbol{X}_n, s, 0)$ is a $G(n, 1/2)$ random graph. With a slight abuse of notation, we write $\Gamma(\boldsymbol{X}_n, s)$ for $\Gamma(\boldsymbol{X}_n, s, 0)$.

We study the *random graph process*

$$\mathbb{G}_{d,p}(\boldsymbol{X}_n) = \left\{ \Gamma(\boldsymbol{X}_n, s, \Phi^{-1}(1-p)) : s \in S^{d-1} \right\} .$$

$\mathbb{G}_{d,p}(\boldsymbol{X}_n)$ is a stationary process of $G(n,p)$ random graphs, indexed by d-dimensional unit vectors. For larger values of d, the process becomes "richer". Our aim is to explore how large the dimension d needs to be for there to exist random directions s for which $\Gamma(\boldsymbol{X}_n, s, \Phi^{-1}(1-p)) \in \mathbb{G}_{d,p}(\boldsymbol{X}_n)$ has different behavior from what is expected from a $G(n,p)$ random graph. Adapting terminology from dynamical percolation [18], we call such directions *exceptional rotations*. More precisely, in analogy with the Vapnik-Chervonenkis theory of studying atypical deviations of averages from their means, our aim is to develop a *VC theory* of random graphs. In particular, we study three fundamental properties of the graphs in the family $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$: the size of the largest clique, the chromatic number, and connectivity. In the first two cases we consider $p = 1/2$ while in the study of connectivity we focus on the case when $p = c \log n / n$ for some constant $c > 0$.

The graph properties we consider are all monotone, so have a critical probability p^* at which they are typically obtained by $G(n,p)$. For example, consider connectivity, and suppose we first place ourselves above the critical probability in $G(n,p)$, e.g., $p = c \log n / n$ for $c > 1$, so that $G(n,p)$ is with high probability connected. Then the question is how large should d be to ensure that for some member graph in the class, the property (connectivity) disappears. There is a threshold *dimension d* for this, and we develop upper and lower bounds for that dimension. Secondly, consider the regime below the critical probability for connectivity in $G(n,p)$, e.g., $p = c \log n / n$ for $c < 1$. In this case, with high probability $G(n,p)$ is not connected, and we ask how large d should be to ensure that for some member graph in the class, the property (connectivity) appears. Again, we develop upper and lower bounds for the threshold dimension d for this.

In all, for each of the three properties considered in this paper, clique number, chromatic number, and connectivity, four theorems are needed, to describe upper and lower bounds for the threshold dimension for exceptional behaviour in the subcritical regime (when the property typically does not obtain) and in the supercritical regime (when the property typically does obtain). In every case, our results reveal a remarkable asymmetry between "upper" and "lower" deviations relative to this threshold.

Our techniques combine some of the essential notions introduced by Vapnik and Chervonenkis (such as shattering, covering, packing, and symmetrization), with elements of high-dimensional random geometry, coupled with sharp estimates for certain random graph parameters.

The model considered in this paper uses subsets of the collection of halfspaces in $\mathbb{R}^d$ to define the random graphs in the collection. A natural variant would be one in which we associate with each edge $\{i,j\}$ a uniformly distributed random vector on the torus $[0,1]^d$, and consider a class parametrized by $s \in [0,1]^d$. Then define the edge set $e(\Gamma(\boldsymbol{X}_n, s, t)) = \{\{i,j\} : \|X_{i,j} - s\| \le t\}$. For general classes of sets of $\mathbb{R}^d$, the complexity of the classes will affect the behaviour of the collection of random graphs in a universal manner. We can define the complexity of a class of graphs indexed in terms of the threshold dimension needed to make certain graph properties appear or disappear in the subcritical and supercritical regimes, respectively. It will be interesting to explore the relationship between the combinatorial geometry of the class and these complexities.

Note that when $d = 1$, $\mathbb{G}_{1,p}(\boldsymbol{X}_n)$ only contains two graphs (when $p = 1/2$, one is the complement of the other), and therefore the class is trivial. On the other extreme, when $d \ge \binom{n}{2}$, with probability one, the collection $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ contains *all* $2^{\binom{n}{2}}$ graphs on n vertices. This follows from the following classical result on the "VC shatter coefficient" of linear half spaces (see, e.g., Schläffli [16], Cover [6]) that determines the number of different graphs in $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ (with probability one).

Lemma 1.1. *Given $N \ge d$ points $x_1, \ldots, x_N \in \mathbb{R}^d$ in general position (i.e., every subset of d points is linearly independent), the number of binary vectors $b \in \{0,1\}^N$ of the form $b = \left(\mathbb{1}_{\{\langle x_i, s \rangle \ge 0\}}\right)_{i \le N}$ for some $s \in S^{d-1}$ equals*

$$C(N, d) = 2 \sum_{k=0}^{d-1} \binom{N-1}{k} .$$

In particular, when $N = d$, all 2^N possible dichotomies of the N points are realizable by some linear half space with the origin on its boundary. In such a case we say that the N points are *shattered* by half spaces.

Notation and Overview. Throughout the paper, log denotes natural logarithm. For a sequence $\{A_n\}$ of events, we say that A_n holds *with high probability* if $\lim_{n \to \infty} \mathbb{P}\{A_n\} = 1$.

The paper is organized as follows. In Section 2 we study the clique number in the case $p = 1/2$. The four parts of Theorem 2.1 establish upper and lower bounds for the critical dimension above which, with high probability, there exist graphs in $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ whose largest clique is significantly larger/smaller than the typical value, which is $\approx 2 \log_2 n - 2 \log_2 \log_2 n$. We show that the critical dimension for which some graphs in $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ have a clique number at least, say, $10 \log_2 n$ is of the order of $\log^2 n / \log \log n$.

In sharp contrast to this, d needs to be at least $n^2 / polylog\, n$ to find a graph in $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ with maximum clique size 3 less than the typical value. We study this functional in Section 3. Theorem 3.1 summarizes the four statements corresponding to upper and lower bounds in the sub-, and super-critical regime. Once again, the two regimes exhibit an important asymmetry. While no graphs in $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ have a chromatic number a constant factor larger than typical unless d is is of the order of $n^2 / polylog\, n$, there exist graphs with a constant factor smaller chromatic number for d near n.

Finally, in Section 4, connectivity properties are examined. To this end, we place ourselves in the regime $p = c \log n / n$ for some constant c. When $c < 1$, a typical graph $G(n,p)$ is disconnected, with high probability, while for $c > 1$ it is connected. In Theorem 4.1 we address both cases. We show that for $c > 1$, the critical dimension above which one finds disconnected graphs among $\mathbb{G}_{d,c\log n/n}(\boldsymbol{X}_n)$ is of the order of $\log n / \log \log n$. (Our upper and lower bounds differ by a factor of 2.) We also show that when $c < 1$, d needs to be at least roughly n^{1-c} in order to find a

connected graph $\mathbb{G}_{d,c\log n/n}(\boldsymbol{X}_n)$. While we conjecture this lower bound to be sharp, we do not have a matching upper bound in this case. However, we are able to show that when d is at least of the order of $n\sqrt{\log n}$, $\mathbb{G}_{d,c\log n/n}(\boldsymbol{X}_n)$ not only contains some connected graphs but with high probability, for any spanning tree, there exists $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}_n, s, t)$ contains that spanning tree. This property holds for even much smaller values of p.

In the Appendix we gather some technical estimates required for the proofs.

2 Clique number

In this section we consider $p = 1/2$ and investigate the extremes of the clique number amongst the graphs $\Gamma(\boldsymbol{X}_n, s)$, $s \in S^{d-1}$. Denote by $cl(\boldsymbol{X}_n, s)$ the size of the largest clique in $\Gamma(\boldsymbol{X}_n, s)$.

The typical behavior of the clique number of a $G(n, 1/2)$ random graph is quite accurately described by Matula's classical theorem [12] that states that for any fixed $s \in S^{d-1}$, for any $\epsilon > 0$,

$$cl(\boldsymbol{X}_n, s) \in \{\lfloor \omega - \epsilon \rfloor, \lfloor \omega + \epsilon \rfloor\}$$

with probability tending to 1, where $\omega = 2\log_2 n - 2\log_2 \log_2 n + 2\log_2 e - 1$.

Here we are interested in understanding the values of d for which graphs with atypical clique number appear. We prove below that while for moderately large values of d some graphs $\Gamma(\boldsymbol{X}_n, s)$ have a significantly larger clique number than ω, one does not find graphs with significantly smaller clique number unless d is nearly quadratic in n.

Observe first that by Lemma 1.1 for any k, if $d \geq \binom{k}{2}$, then, with probability one, $cl(\boldsymbol{X}_n, s) \geq k$ for some $s \in S^{d-1}$. (Just fix any set of k vertices; all 2^k graphs on these vertices is present for some s, including the complete graph.) For example, when $d \sim (9/2)(\log_2 n)^2$, $cl(\boldsymbol{X}_n, s) \geq 3\log_2 n$ for some $s \in S^{d-1}$, a quite atypical behavior. In fact, with a more careful argument we show below that when d is a sufficiently large constant multiple of $(\log n)^2/\log \log n$, then, with high probability, there exists $s \in S^{d-1}$ such that $cl(\boldsymbol{X}_n, s) \geq 3\log_2 n$. We also show that no such s exists for $d = o((\log n)^2/\log \log n)$. Perhaps more surprisingly, clique numbers significantly smaller than the typical value only appear for huge values of d. The next theorem shows the surprising fact that in order to have that for some $s \in S^{d-1}$, $cl(\boldsymbol{X}_n, s) < \omega - 3$, the dimension needs to be $n^{2-o(1)}$. (Recall that for $d = \binom{n}{2}$ the point set $\boldsymbol{X}_n$ is shattered and one even has $cl(\boldsymbol{X}_n, s) = 1$ for some s. Our findings on the clique number are summarized in the following theorem.

Theorem 2.1. (CLIQUE NUMBER.) *If $cl(\boldsymbol{X}_n, s)$ denotes the clique number of $\Gamma(\boldsymbol{X}_n, s)$, then, with high probability the following hold:*

(i) (SUBCRITICAL; NECESSARY.) *If $d = o(n^2/(\log n)^9)$, then for all $s \in S^{d-1}$, $cl(\boldsymbol{X}_n, s) > \omega - 3$.*

(ii) (SUBCRITICAL; SUFFICIENT.) *If $d \geq \binom{n}{2}$, then there exists $s \in S^{d-1}$ such that $cl(\boldsymbol{X}_n, s) = 1$.*

(iii) (SUPERCRITICAL; NECESSARY.) *For any $c > 2$ there exists $c' > 0$ such that if $d \leq c'\log^2 n/\log \log n$, then for all $s \in S^{d-1}$, we have $cl(\boldsymbol{X}_n, s) \leq c\log_2 n$.*

(iv) (SUPERCRITICAL; SUFFICIENT.) *For any $c > 2$ and $c' > c^2/(2\log 2)$, if $d \geq c'\log^2 n/\log \log n$, then there exists $s \in S^{d-1}$ such that $cl(\boldsymbol{X}_n, s) \geq c\log_2 n$.*

The event described in (ii) holds with probability one for all n.

Proof. To prove part (i), let $k = \lfloor \omega - 3 \rfloor$ and let $N_k(s)$ denote the number of cliques of size k in $\Gamma(\boldsymbol{X}_n, s)$. Let $\eta \in (0, 1]$ and let $\mathcal{C}_\eta$ be a minimal η-cover of S^{d-1}. Then

$$\mathbb{P}\left\{\exists s \in S^{d-1} : N_k(s) = 0\right\}$$
$$= \mathbb{P}\left\{\exists s' \in \mathcal{C}_\eta \text{ and } \exists s \in S^{d-1} : \|s - s'\| \leq \eta : N_k(s) = 0\right\}$$
$$\leq |\mathcal{C}_\eta|\mathbb{P}\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N_k(s) = 0\right\}$$

where $s_0 = (1, 0, \ldots, 0)$ and the last inequality follows from the union bound. Consider the graph $\Gamma(\boldsymbol{X}_n, s_0, -\eta\sqrt{1 - \eta^2/2})$ in which vertex i and vertex j are connected if and only if the first component of $X_{i,j}$ is at least $-\eta\sqrt{1 - \eta^2/2}$

The proof of Lemma 5.3 implies that the event $\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N_k(s) = 0\right\}$ is included in the event that $\Gamma(\boldsymbol{X}_n, s_0, -\eta\sqrt{1 - \eta^2/2})$ does not have any clique of size k. By Lemma 5.3, the probability of this is bounded by the probability that an Erdős-Rényi random graph $G(n, 1/2 - \alpha_n)$ does not have any clique of size k where $\alpha_n = \frac{\eta\sqrt{d}}{\sqrt{2\pi}}$. If we choose (say) $\eta = 1/n^2$ then for $d \leq n^2$ we have $\alpha_n \leq 1/n$ and therefore, by Lemma 5.7 below,

$$\mathbb{P}\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N_k(s) = 0\right\} \leq \exp\left(\frac{-C'n^2}{(\log_2 n)^8}\right)$$

for some numerical constant C'. Thus, using Lemma 5.1,

$$\mathbb{P}\left\{\exists s \in S^{d-1} : N_k(s) = 0\right\} \leq (4n^2)^d \exp\left(\frac{-C'n^2}{(\log_2 n)^8}\right) = o(1)$$

whenever $d = o(n^2/(\log n)^9)$.

Part (ii) follows from the simple fact that, by Lemma 1.1, with $d = \binom{n}{2}$ even the empty graph appears among the $\Gamma(\boldsymbol{X}_n, s)$.

The proof of part (iii) proceeds similarly to that of part (i). Let $k = c\log_2 n$. Then

$$\mathbb{P}\left\{\exists s \in S^{d-1} : N_k(s) \geq 1\right\}$$
$$\leq |\mathcal{C}_\eta|\mathbb{P}\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N_k(s) \geq 1\right\}.$$

Similarly to the argument of (i), we note that the event $\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N_k(s) \geq 1\right\}$ is included in the event that $\Gamma(\boldsymbol{X}_n, s_0, -\eta\sqrt{1 - \eta^2/2})$ has a clique of size k, which is bounded by the probability that an Erdős-Rényi random graph $G(n, 1/2 + \alpha_n)$ has a clique of size k where $\alpha_n = \frac{\eta\sqrt{d}}{\sqrt{2\pi}}$. Denoting $p = 1/2 + \alpha_n$, this probability is bounded by $\binom{n}{k}p^{\binom{k}{2}} \leq \left(np^{k/2}\right)^k$. We may choose $\eta = 4/d$. Then, for d sufficiently large, $\alpha_n \leq (c/2 - 1)\log 2$ and, using Lemma 5.1, we have

$$\mathbb{P}\left\{\exists s \in S^{d-1} : N_k(s) \geq 1\right\} \leq (4/\eta)^d \left(np^{(c/2)\log_2 n}\right)^{c\log_2 n}$$
$$\leq e^{d\log d}\left(n^{1 + (c/2)\log_2(1/2 + \alpha_n)}\right)^{c\log_2 n}$$
$$\leq e^{d\log d}\left(n^{1 - c/2 + c\alpha_n/\log 2}\right)^{c\log_2 n}$$
$$\leq e^{d\log d}n^{(1 - c/2)c(\log_2 n)/2}$$
$$= e^{d\log d - (c-2)c(\log_2 n)^2\log 2/4},$$

5

and the statement follows.

It remains to prove part (iv). The proof relies on the second moment method. Let $c > 2$, $c' > c^2/(2 \log 2)$, and assume that $d \geq c' \log^2 n / \log \log n$. Let K be a constant satisfying $K > 2/\sqrt{c'}$ and define $\theta = K\sqrt{\log \log n}/\log n$. Let A be a subset of S^{d-1} of cardinality at least $(d/16)\theta^{-(d-1)}$ such that for all distinct pairs $s, s' \in A$, we have $\langle s, s' \rangle \geq \cos(\theta)$. Such a set exists by Lemma 5.2. Also, let $\mathcal{C}$ be the family of all subsets of $[n]$ of cardinality $k = \lfloor c \log_2 n \rfloor$. For $s \in A$ and $\gamma \in \mathcal{C}$, denote by $Z_{s,\gamma}$ the indicator that all edges between vertices in γ are present in the graph $\Gamma(\boldsymbol{X}_n, s)$. Our aim is to show that $\lim_{n\to\infty} \mathbb{P}\{Z > 0\} = 1$ where

$$Z = \sum_{s \in A} \sum_{\gamma \in \mathcal{C}} Z_{s,\gamma} \ .$$

To this end, by the second moment method (see, e.g., [1]), it suffices to prove that $\mathbb{E}Z \to \infty$ and that $\mathbb{E}[Z^2] = (\mathbb{E}Z)^2(1 + o(1))$.

To bound $\mathbb{E}Z$ note that

$$
\begin{aligned}
\mathbb{E}Z &= |A|\binom{n}{k}\mathbb{E}Z_{s,\gamma} \\
&\geq (d/16)\theta^{-(d-1)}\binom{n}{k}2^{-\binom{k}{2}} \\
&= \exp\left((\log n)^2 \left(c' - \frac{c^2}{2 \log 2} + \frac{c}{\log 2} + o(1)\right)\right) \to \infty \ .
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\mathbb{E}[Z^2] &= \sum_{s,s' \in A} \sum_{\gamma,\gamma' \in \mathcal{C}} \mathbb{E}[Z_{s,\gamma}Z_{s',\gamma'}] \\
&= \sum_{s,s':s\neq s' \in A} \sum_{\gamma,\gamma' \in \mathcal{C}:|\gamma\cap\gamma'|\leq 1} \mathbb{E}[Z_{s,\gamma}Z_{s',\gamma'}] + \sum_{s \in A} \sum_{\gamma,\gamma' \in \mathcal{C}} \mathbb{E}[Z_{s,\gamma}Z_{s,\gamma'}] \\
&\quad + \sum_{s,s':s\neq s' \in A} \sum_{\gamma,\gamma' \in \mathcal{C}:|\gamma\cap\gamma'|\geq 2} \mathbb{E}[Z_{s,\gamma}Z_{s',\gamma'}] \\
&\stackrel{\text{def}}{=} I + II + III \ .
\end{aligned}
$$

For the first term note that if γ and γ' intersect in at most one vertex then $Z_{s,\gamma}$ and $Z_{s',\gamma'}$ are independent and therefore

$$I = \sum_{s,s':s\neq s' \in A} \sum_{\gamma,\gamma' \in \mathcal{C}:|\gamma\cap\gamma'|\leq 1} \mathbb{E}Z_{s,\gamma}\mathbb{E}Z_{s',\gamma'} \leq (\mathbb{E}Z)^2 \ .$$

Hence, it suffices to prove that $II + III = o((\mathbb{E}Z)^2)$. To deal with II, we have

$$
\begin{aligned}
\frac{II}{(\mathbb{E}Z)^2} &= \frac{1}{|A| \cdot \binom{n}{k}} \sum_{\ell=0}^{k} 2^{\binom{\ell}{2}} \binom{n-k}{k-\ell} \binom{k}{\ell} \\
&\leq \frac{1}{|A|} \sum_{\ell=0}^{k} 2^{\binom{\ell}{2}} \frac{k^{2\ell}}{(n-2k)^\ell \ell!} \\
&\leq \frac{1}{|A|} 2^{\binom{\ell}{2}} \sum_{\ell=0}^{\infty} \left(\frac{k^2}{n-2k}\right)^\ell \frac{1}{\ell!} \\
&= \exp\left(-(\log n)^2 \left(c' + o(1) - c^2/(2\log 2)\right)\right) \to 0 .
\end{aligned}
$$

We now take care of III. To this end, we bound

$$
\max_{\substack{s,s' \in A : s \neq s' \\ \gamma,\gamma' : |\gamma \cap \gamma'| = \ell}} \mathbb{E}[Z_{s,\gamma} Z_{s',\gamma'}]
$$

by

$$
2^{\binom{\ell}{2} - 2\binom{k}{2} + 1} \mathbb{P}\left\{ \left\langle \frac{N}{\|N\|}, s_0 \right\rangle \geq \sin(\theta/2) \right\}^{\binom{\ell}{2}} ,
$$

where N is a standard normal vector in $\mathbb{R}^d$. To see this, note that $2\binom{k}{2} - \binom{\ell}{2}$ edges of the two cliques occur independently, each with probability $1/2$. The remaining $\binom{\ell}{2}$ edges must be in both $\Gamma(\boldsymbol{X}_n, s)$ and $\Gamma(\boldsymbol{X}_n, s')$. A moment of thought reveals that this probability is bounded by the probability that the angle between a random normal vector and a fixed unit vector (say s_0) is less than $\pi/2 - \theta/2$. This probability may be bounded as

$$
\begin{aligned}
\mathbb{P}\left\{ \langle N/\|N\|, s_0 \rangle \geq \sin(\theta/2) \right\} &= \frac{1}{2} \mathbb{P}\left\{ B \geq \sin^2(\theta/2) \right\} \\
&\qquad \text{(where } B \text{ is a Beta}(1/2, (d-1)/2) \text{ random variable)} \\
&\leq \frac{\mathbb{E}B}{2\sin^2(\theta/2)} \\
&= \frac{1}{2d\sin^2(\theta/2)} \\
&= \frac{2 + o(1)}{d\theta^2} = \frac{2 + o(1)}{c' K^2} .
\end{aligned}
$$

Via the same counting argument used in handling II, we have

$$
\frac{III}{(\mathbb{E}Z)^2} \leq \sum_{\ell=2}^{k} 2^{\binom{\ell}{2}} \left(\frac{2 + o(1)}{c' K^2}\right)^{\binom{\ell}{2}} \left(\frac{k^2}{n-2k}\right)^\ell \frac{1}{\ell!} .
$$

Since $c' K^2 > 4$, we have, for n large enough,

$$
\frac{III}{(\mathbb{E}Z)^2} \leq \sum_{\ell=2}^{k} \left(\frac{k^2}{n-2k}\right)^\ell \frac{1}{\ell!} = O\left(\frac{(\log n)^2}{n^2}\right)
$$

as required. This concludes the proof of the theorem. □

We conclude the section by remarking that the above proof extends straightforwardly to $G(n, p)$ for any constant $p \in (0, 1)$.

3 Chromatic number

A proper coloring of vertices of a graph assigns a color to each vertex such that no pair of vertices joined by an edge share the same color. The *chromatic number* $\chi(G)$ of a graph G is the smallest number of colors for which a proper coloring of the graph exists.

Here we study the fluctuations of the chromatic numbers $\chi(\Gamma(\boldsymbol{X}_n, s))$ from its typical behavior as $s \in S^{d-1}$. Once again, for simplicity of the presentation, we consider $p = 1/2$. The arguments extend easily to other (constant) values of p.

For a fixed s, a celebrated result of Bollobás [3] implies that

$$\frac{n}{2 \log_2 n} \leq \chi(\Gamma(\boldsymbol{X}_n, s)) \leq \frac{n}{2 \log_2 n}(1 + o(1))$$

with high probability.

In this section we derive estimates for the value of the dimension d for which there exist random graphs in the collection $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ whose chromatic number differs substantially (i.e., by a constant factor) from that of a typical $G(n, 1/2)$ graph. Similar to the case of the clique number studied in Section 2, we find that upper and lower deviations exhibit a different behavior—though in a less dramatic way. With high probability, one does not see a graph with a clique number larger than $(1 + \epsilon)n/(2 \log_2 n)$ unless d is at least $n^2/polylog\, n$. On the other hand, when d is roughly linear in n, there are graphs is $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$ with chromatic number at most $(1 - \epsilon)n/(2 \log_2 n)$. Below we make these statements rigorous and also show that they are essentially tight.

Theorem 3.1. (CHROMATIC NUMBER.) *Let $\epsilon \in (0, 1/2)$. If $\chi(\Gamma(\boldsymbol{X}_n, s))$ denotes the chromatic number of $\Gamma(\boldsymbol{X}_n, s)$, then, with high probability the following hold:*

(i) (SUBCRITICAL; NECESSARY.) *If $d = o(n/(\log n)^3)$, then for all $s \in S^{d-1}$, $\chi(\Gamma(\boldsymbol{X}_n, s)) \geq (1 - \epsilon)n/(2 \log_2 n)$.*

(ii) (SUBCRITICAL; SUFFICIENT.) *If $d \geq 2n \log_2 n/(1 - 2\epsilon)$, then there exists $s \in S^{d-1}$ such that $\chi(\Gamma(\boldsymbol{X}_n, s)) \leq (1 - \epsilon)n/(2 \log_2 n)$.*

(iii) (SUPERCRITICAL; NECESSARY.) *If $d = o(n^2/(\log n)^6)$, then for all $s \in S^{d-1}$, $\chi(\Gamma(\boldsymbol{X}_n, s)) \leq (1 + \epsilon)n/(2 \log_2 n)$.*

(iv) (SUPERCRITICAL; SUFFICIENT.) *If $d \geq .5\left[(1 + \epsilon)n/(2 \log_2 n)\right]^2$, then there exists $s \in S^{d-1}$ such that $\chi(\Gamma(\boldsymbol{X}_n, s)) \geq (1 + \epsilon)n/(2 \log_2 n)$.*

Part (i) of Theorem 3.1 follows from the following "uniform concentration" argument.

Proposition 3.2. *If $d = o(n/(\log n)^3)$, we have*

$$\sup_{s \in S^{d-1}} \left| \chi(\Gamma(\boldsymbol{X}_n, s)) - \frac{n}{2 \log_2 n} \right| = o_p\left(\frac{n}{\log_2 n}\right) ,$$

Proof. A classical result of Shamir and Spencer[17] shows that for any fixed $s \in S^{d-1}$,

$$|\chi(\Gamma(\boldsymbol{X}_n, s)) - \mathbb{E}\left(\chi(\Gamma(\boldsymbol{X}_n, s))\right)| = O_p(n^{1/2}) .$$

In fact, one may easily combine the above-mentioned results of Bollobás and Shamir and Spencer to obtain that
$$\frac{\mathbb{E}\chi(\Gamma(\boldsymbol{X}_n, s))}{n/(2\log_2 n)} \to 1 \ .$$

The proof of the proposition is based on combining the Shamir-Spencer concentration argument with Vapnik-Chervonenkis-style symmetrization.

For each $s \in S^{d-1}$ and $i = 2, \ldots, n$, define $Y_{i,s} = (\mathbb{1}_{\{\langle X_{i,j}, s\rangle \geq 0\}})_{j=1,\ldots,i-1} \in \{0,1\}^{i-1}$ as the collection of indicators of edges connecting vertex i smaller-labeled vertices in $\Gamma(\boldsymbol{X}_n, s)$. As Shamir and Spencer, we consider the chromatic number $\Gamma(\boldsymbol{X}_n, s)$ as a function of these variables and define the function $f : \prod_{i=2}^n \{0,1\}^{i-1} \to \mathbb{N}$ by

$$f(Y_{2,s}, \ldots, Y_{n,s}) = \chi(\Gamma(\boldsymbol{X}_n, s)) \ .$$

By Markov's inequality, it suffices to show that

$$\mathbb{E}\left[\sup_{s \in S^{d-1}} |f(Y_{2,s}, \ldots, Y_{n,s}) - \mathbb{E}f(Y_{2,s}, \ldots, Y_{n,s})|\right] = o\left(\frac{n}{\log n}\right) \ .$$

Let $\boldsymbol{X}'_n = (X'_{i,j})_{1 \leq i < j \leq n}$ be an independent copy of $\boldsymbol{X}_n$. Denote by $\mathbb{E}'$ conditional expectation given $\boldsymbol{X}_n$. We write $Y'_{i,s} = (\mathbb{1}_{\{\langle X'_{i,j}, s\rangle \geq 0\}})_{j=1,\ldots,i-1} \in \{0,1\}^{i-1}$.

Also introduce random "swap operators" $\epsilon_2, \ldots, \epsilon_n$ defined by

$$\epsilon_i(Y_{i,s}, Y'_{i,s}) = \begin{cases} Y_{i,s} & \text{with probability } 1/2 \\ Y'_{i,s} & \text{with probability } 1/2 \end{cases}$$

where the ϵ_i are independent of each other and of everything else.

$$\mathbb{E}\left[\sup_{s \in S^{d-1}} |f(Y_{2,s}, \ldots, Y_{n,s}) - \mathbb{E}f(Y_{2,s}, \ldots, Y_{n,s})|\right]$$

$$= \mathbb{E}\left[\sup_{s \in S^{d-1}} \left|\mathbb{E}'\left(f(Y_{2,s}, \ldots, Y_{n,s}) - f(Y'_{2,s}, \ldots, Y'_{n,s})\right)\right|\right]$$

$$\leq \mathbb{E}\left[\sup_{s \in S^{d-1}} \left|f(Y_{2,s}, \ldots, Y_{n,s}) - f(Y'_{2,s}, \ldots, Y'_{n,s})\right|\right]$$

$$= \mathbb{E}\left[\sup_{s \in S^{d-1}} \left|f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s})) - f(\epsilon_2(Y'_{2,s}, Y_{2,s}), \ldots, \epsilon_n(Y'_{n,s}, Y_{n,s}))\right|\right] \ .$$

Introduce now the expectation operator $\mathbb{E}_\epsilon$ that computes expectation with respect to the random swaps only. Then we can further bound the expectation above by

$$2\mathbb{E}\mathbb{E}_\epsilon\left[\sup_{s \in S^{d-1}} \left|f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s})) - \mathbb{E}_\epsilon f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s}))\right|\right] \ .$$

Next we bound the inner expectation. Note that for fixed $\boldsymbol{X}_n, \boldsymbol{X}'_n$, by Lemma 1.1, there are at most n^{2d} different dichotomies of the $2\binom{n}{2}$ points in $\boldsymbol{X}_n \cup \boldsymbol{X}'_n$ by hyperplanes including the origin and therefore there are not more than n^{2d} random variables of the form $f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s}))$

9

as s varies over S^{d-1}. On the other hand, for any fixed s, the value of $f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s}))$ can change by at most 1 if one flips the value of one of the $\epsilon_i(Y_{i,s}, Y'_{i,s})$ $(i = 2, \ldots, n)$, since such a flip amounts to changing the edges incident to vertex i and therefore can change the value of the chromatic number by at most one. Thus, by the bounded differences inequality (see, e.g., [5, Section 6.1]), for all $s \in S^{d-1}$ and $\lambda > 0$,

$$\mathbb{E}_\epsilon \left[\exp \left(\lambda(f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s})) - \mathbb{E}_\epsilon f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s})))) \right] \right.$$
$$\leq \exp\left(\frac{(n-1)\lambda^2}{2} \right).$$

Therefore, by a standard maximal inequality for sub-Gaussian random variables ([5, Section 2.5]),

$$\mathbb{E}_\epsilon \left[\sup_{s \in S^{d-1}} \left| f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s})) - \mathbb{E}_\epsilon f(\epsilon_2(Y_{2,s}, Y'_{2,s}), \ldots, \epsilon_n(Y_{n,s}, Y'_{n,s})) \right| \right]$$
$$\leq \sqrt{4(n-1)d \log n}.$$

Since the upper bound is $o(n/\log n)$ for $d = o(n/\log^3 n)$, the result follows. □

Parts (ii) and (iv) of Theorem 3.1 follow from the next, straingforward proposition by setting $k = \lfloor (1-\epsilon)n/(2\log_2 n) \rfloor$ and $k' = \lceil (1+\epsilon)n/(2\log_2 n) \rceil$.

Proposition 3.3. *Let $k, k' \leq n$ be positive integers. If $d \geq k\binom{\lceil n/k \rceil}{2}$, then, with probability one, there exists $s \in S^{d-1}$ such that $\chi(\Gamma(\boldsymbol{X}_n, s)) \leq k$. On the other hand, if $d \geq \binom{k'}{2}$, then, with probability one, there exists $s \in S^{d-1}$ such that $\chi(\Gamma(\boldsymbol{X}_n, s)) \geq k'$.*

Proof. Partition the vertex set $[n]$ into k disjoint sets of size at most $\lceil n/k \rceil$ each. If for some $s \in S^{d-1}$ each of these sets is an independent set (i.e., contain no edge joining two vertices within the set) in $\Gamma(\boldsymbol{X}_n, s)$, then the graph $\Gamma(\boldsymbol{X}_n, s)$ is clearly properly colorable with k colors. Let A be the set of pairs of vertices (i, j) such that i and j belong to the same set of the partition. By Lemma 1.1, if $d \geq k\binom{\lceil n/k \rceil}{2} \geq |A|$, the set of points $\{X_{i,j} : (i, j) \in A\}$ is shattered by half spaces. In particular, there exists an $s \in S^{d-1}$ such that $\langle X_{i,j}, s \rangle < 0$ for all $(i, j) \in A$ and therefore $\Gamma(\boldsymbol{X}_n, s)$ has no edge between any two vertices in the same set. The first statement follows.

To prove the second statement, simply notice that is a graph has a clique of size k then its chromatic number at least k. But if $d \geq \binom{k}{2}$, then, by Lemma 1.1, for some $s \in S^{d-1}$, the vertex set $\{1, \ldots, k\}$ forms a clique. □

It remains to prove Part (iii) of Theorem 3.1. To this end, we combine the covering argument used in parts (i) and (iii) of Theorem 2.1 with a result of Alon and Sudakov [2] (see Proposition 5.8 below) that bounds the "resilience" of the chromatic number of a random graph.

Let $\mathcal{C}_\eta$ be a minimal η-cover of S^{d-1} where we take $\eta = c\epsilon^2/(\sqrt{d}\log^2 n)$ for a sufficiently small positive constant c. Then

$$\mathbb{P}\left\{ \exists s \in S^{d-1} : \chi(\Gamma(\boldsymbol{X}_n, s)) > (1+\epsilon)\frac{n}{2\log_2 n} \right\}$$
$$\leq |\mathcal{C}_\eta| \mathbb{P}\left\{ \exists s \in S^{d-1} : \|s - s_0\| \leq \eta : \chi(\Gamma(\boldsymbol{X}_n, s)) > (1+\epsilon)\frac{n}{2\log_2 n} \right\}$$

10

where $s_0 = (1, 0, \ldots, 0)$. By the argument used in the proof of parts (i) and (iii) of Theorem 2.1,

$$\bigcup_{s \in S^{d-1} : \|s - s_0\| \leq \eta} \Gamma(\boldsymbol{X}_n, s) \subset \Gamma(\boldsymbol{X}_n, s_0) \cup E$$

where E is a set of $\mathrm{Bin}(\binom{n}{2}, \alpha_n)$ edges where, $\alpha_n = \frac{\eta \sqrt{d}}{\sqrt{2\pi}}$. By our choice of η, we have $\alpha_n \leq c_2 \epsilon^2 n^2 / (\log_2 n)^2$ where c_2 is the constant appearing in Proposition 5.8. Thus, by the Chernoff bound,

$$\mathbb{P}\left\{ |E| > \frac{c_2 \epsilon^2 n^2}{(\log_2 n)^2} \right\} \leq \exp\left(-\frac{c_2 (\log 2 - 1/2) \epsilon^2 n^2}{(\log_2 n)^2} \right).$$

Hence, by Proposition 5.8,

$$\mathbb{P}\left\{ \exists s \in S^{d-1} : \|s - s_0\| \leq \eta : \chi(\Gamma(\boldsymbol{X}_n, s)) > (1 + \epsilon) \frac{n}{2 \log_2 n} \right\}$$

$$\leq \exp\left(-\frac{c_2 (\log 2 - 1/2) \epsilon^2 n^2}{(\log_2 n)^2} \right) + \exp\left(-\frac{c_1 n^2}{(\log_2 n)^4} \right).$$

Combining this bound with Lemma 5.1 implies the statement. $\qquad \square$

4 Connectivity

In this section we study connectivity of the random graphs in $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$. It is well known since the pioneering work of Erdős and Rényi [7] that the threshold for connectivity for a $G(n, p)$ random graph is when $p = c \log n/n$. For $c < 1$, the graph is disconnected and for $c > 1$ it is connected, with high probability. In this section we investigate both regimes. In particular, for $c > 1$ we establish lower and upper bounds for the smallest dimension d such that some graph in $\mathbb{G}_{d,c \log n/n}(\boldsymbol{X}_n)$ is disconnected. We prove that this value of d is of the order of $(c - 1) \log n / \log \log n$. For the regime $c < 1$ we also establish lower and upper bounds for the smallest dimension d such that some graph in $\mathbb{G}_{d,c \log n/n}(\boldsymbol{X}_n)$ is connected. As in the case of the clique number and chromatic number, here as well we observe a large degree of asymmetry. In order to witness some connected graphs in $\mathbb{G}_{d,c \log n/n}(\boldsymbol{X}_n)$, the dimension d has to be at least of the order of n^{1-c}. While we suspect that this bound is essentially tight, we do not have a matching upper bound. However, we are able to show that when d is of the order of $n \log n$, the family $\mathbb{G}_{d,c \log n/n}(\boldsymbol{X}_n)$ not only contains connected graphs, but also, with high probability, for every spanning tree of the vertices $[n]$, there exists an $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}_n, s, t)$ contains the spanning tree. (Recall that t is such that $p = 1 - \Phi(t)$.)

Theorem 4.1. (CONNECTIVITY.) *Assume $p = c \log n/n$ and let $t = \Phi^{-1}(1 - p)$. Then with high probability the following hold:*

(i) (SUBCRITICAL; NECESSARY.) *If $c < 1$ then for any $\epsilon \in (0, c)$, if $d = O(n^{1-c-\epsilon})$, then for all $s \in S^{d-1}$, $\Gamma(\boldsymbol{X}_n, s, t)$ is disconnected.*

(ii) (SUBCRITICAL; SUFFICIENT.) *There exists an absolute constant C such that if $d \geq C n \sqrt{\log n}$, then there exists an $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}_n, s, t)$ is connected.*

(iii) (SUPERCRITICAL; NECESSARY.) *If $c > 1$ then for any $\epsilon > 0$, if $d \leq (1-\epsilon)(c-1) \log n / \log \log n$, then for all $s \in S^{d-1}$, $\Gamma(\boldsymbol{X}_n, s, t)$ is connected.*

11

(iv) (SUPERCRITICAL; SUFFICIENT.) *If $c > 1$ then for any $\epsilon > 0$, if $d \geq (2+\epsilon)(c-1)\log n/\log\log n$, then for some $s \in S^{d-1}$, $\Gamma(\boldsymbol{X}_n, s, t)$ is disconnected.*

Proof of Theorem 4.1, part (i).

To prove part (i), we show that when $d = O(n^{1-c-\epsilon})$, with high probability, all graphs $\Gamma(\boldsymbol{X}_n, s, t)$ contain at least one isolated point. The proof of this is based on a covering argument similar those used in parts of Theorems 2.1 and 3.1, combined with a sharp estimate for the probability that $G(n, c\log n/n)$ has no isolated vertex. This estimate, given in Lemma 5.9 below, is proved by an elegant argument of O'Connell [13].

Let $\eta \in (0, 1]$ to be specified below and let $\mathcal{C}_\eta$ be a minimal η-cover of S^{d-1}. If $N(s)$ denotes the number of isolated vertices (i.e., vertices of degree 0) in $\Gamma(\boldsymbol{X}_n, s, t)$, then

$$\mathbb{P}\left\{\exists s \in S^{d-1} : \Gamma(\boldsymbol{X}_n, s, t) \text{ is connected}\right\}$$
$$\leq \quad \mathbb{P}\left\{\exists s \in S^{d-1} : N(s) = 0\right\}$$
$$\leq \quad |\mathcal{C}_\eta|\mathbb{P}\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N(s) = 0\right\}$$

where $s_0 = (1, 0, \ldots, 0)$. It follows by the first half of Lemma 5.4 that there exists a constant $\kappa > 0$ such that if $\eta = \kappa\epsilon/(t\sqrt{d})$, then

$$\mathbb{P}\left\{\exists s \in S^{d-1} : \|s - s_0\| \leq \eta : N_k(s) = 0\right\} \leq \mathbb{P}\{N = 0\}$$

where N is the number of isolated vertices in a $G(n, (c + \epsilon/2)\log n/n)$ random graph. By Lemma 5.9, for n sufficiently large, this is at most $\exp(-n^{-(1-c-\epsilon/2)}/3)$. Bounding $|\mathcal{C}_\eta|$ by Lemma 5.1 and substituting the chosen value of η proves part (i).

Proof of Theorem 4.1, part (ii).

Part (ii) of Theorem 4.1 follows from a significantly more general statement. Based on a geometrical argument, we show that for any positive integer k, if d is at least a sufficiently large constant multiple of $k\Phi^{-1}(1-p)$, then with high probability, k independent standard normal vectors in $\mathbb{R}^d$ are shattered by half spaces of the form $\{x : \langle x, s\rangle \geq t\}$. In particular, by taking $k = n - 1$ and considering the normal vectors $X_{i,j}$ corresponding to the edges of any fixed spanning tree, one finds an $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}_n, s, t)$ contains all edges of the spanning tree, making the graph connected. Note that if $d \geq Cn\sqrt{\alpha\log n}$ then the same statement holds whenever $p = n^{-\alpha}$ regardless of how large α is. Thus, for $d \gg n\sqrt{\log n}$, some $\Gamma(\boldsymbol{X}_n, s, t)$ are connected, even though for a typical s, the graph is empty with high probability.

Fix a set E of edges of the complete graph K_n. We say that $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$ shatters E if $\{e(G) : G \in \mathbb{G}_{d,p}(\boldsymbol{X}_n)\}$ shatters E (where $e(G)$ denotes the set of edges of a graph G). In other words, $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$ shatters E if for all $F \subset E$ there is $G \in \mathbb{G}_{d,p}(\boldsymbol{X}_n)$ such that $e(G) \cap E = F$.

Proposition 4.2. *Fix $n \in \mathbb{N}$, $k \in \{1, 2, \ldots, \binom{n}{2}\}$, and a set $E = \{e_1, \ldots, e_k\}$ of edges of the complete graph K_n. There exist universal constants $b, c > 0$ such that for $d \geq (4/c) \cdot k \cdot \Phi^{-1}(1-p)$ we have*

$$\mathbb{P}\left(\mathbb{G}_{d,p}(\boldsymbol{X}_n) \text{ shatters } E\right) \geq 1 - e^{-bd} .$$

Proof. Given points $x_1, \ldots, x_k$ in $\mathbb{R}^d$, the *affine span* of $x_1, \ldots, x_k$ is the set $\{\sum_{i=1}^{k} c_i X_i : \sum_{i=1}^{k} c_i = 1\}$. Fix $E = \{e_1, \ldots, e_k\} \in \mathcal{S}_k$ and let P_E be the affine span of $X_{e_1}, \ldots, X_{e_k}$. Also, let $t = \Phi^{-1}(1-p)$.

First suppose that $\min\{\|y\| : y \in P_E\} > t$. Then we may shatter E as follows. First, almost surely, P_E is a $(k-1)$-dimensional affine subspace in $\mathbb{R}^d$. Assuming this occurs, then E is shattered by halfspaces in P_E: in other words, for any $F \subset E$ there is a $(k-2)$-dimensional subspace H contained within P_E such that F and $E \setminus F$ lie on opposite sides of H in P_E (i.e., in different connected components of $P_E \setminus H$).

Fix $F \subset E$ and $H \subset P_E$ as in the preceding paragraph. Then let K be a $(d-1)$-dimensional hyperplane tangent to $tS^{d-1} = \{x \in \mathbb{R}^d : \|x\| = t\}$, intersecting P_E only at H, and separating the origin from F. In other words, K is such that $K \cap P_E = H$ and $|K \cap tS^{d-1}| = 1$, and also such that 0 and F lie on opposite sides of K of $\mathbb{R}^d \setminus K$. Since P_E has dimension $k - 1 < d - 2$, such a hyperplane K exists. Since F and $E \setminus F$ lie on opposite sides of H, we also obtain that 0 and $E \setminus F$ lie on the *same* side of K.

Let $s \in S^{d-1}$ be such that $ts \in K$. Then for $e \in F$ we have $\langle X_e, s \rangle > t$, and for $e \in E \setminus F$ we have $\langle X_e, s \rangle < t$. It follows that $E \cap \Gamma(\boldsymbol{X}, s, t) = F$. Since $F \subset E$ was arbitrary, this implies that

$$\mathbb{P}(\mathbb{G}_{d,p}(\boldsymbol{X}_n) \text{ shatters } E) \geq \mathbb{P}(\min\{\|y\| : y \in P_E\} > \Phi^{-1}(1-p)),$$

In light of the assumption that $d \geq (4/c) \cdot k \cdot \Phi^{-1}(1-p)$, the proposition is then immediate from Lemma 4.3 below. $\qquad\square$

The key element of the proof of Proposition 4.2 is that the affine span of $k \leq 4d$ independent standard normal vectors in $\mathbb{R}^d$ is at least at distance of the order of d/k from the origin. This is made precise in the following lemma whose proof crucially uses a sharp estimate for the smallest singular value of a $d \times k$ Wishart matrix, due to Rudelson and Vershynin [15].

Lemma 4.3. *There exist universal constants $b, c > 0$ such that the following holds. Let $N_1, \ldots, N_k$ be independent standard normal vectors in $\mathbb{R}^d$, let P be the affine span of $N_1, \ldots, N_k$, and let $D = \min\{\|y\| : y \in P\}$. Then whenever $d \geq 4k$, we have $\mathbb{P}(D \leq cd/4k) < 2e^{-bd}$.*

Proof. We use the notation $\boldsymbol{y} = (y_1, \ldots, y_k)$. We have

$$
\begin{aligned}
D &= \min_{\boldsymbol{y}:\sum y_i = 1} \left\| \sum_{i=1}^{k} y_i N_i \right\|^2 \\
&= \min_{\boldsymbol{y}:\sum y_i = 1} |\boldsymbol{y}|^2 \left\| \sum_{i=1}^{k} \frac{y_i}{\|\boldsymbol{y}\|} N_i \right\|^2 \\
&\geq \min_{\boldsymbol{y}:\sum y_i = 1} \frac{1}{k} \left\| \sum_{i=1}^{k} \frac{y_i}{\|\boldsymbol{y}\|} N_i \right\|^2 \\
&\geq \frac{1}{k} \min_{\boldsymbol{y}:|\boldsymbol{y}|^2 = 1} \left\| \sum_{i=1}^{k} y_i N_i \right\|^2,
\end{aligned}
$$

where the first inequality holds because if $\sum_{i=1}^{k} y_i = 1$ then $\|\boldsymbol{y}\|^2 \geq k^{-1}$ and the second by noting that the vector $(y_i/\|\boldsymbol{y}\|, 1 \leq i \leq k)$ has 2-norm 1.

Let $\boldsymbol{N}$ be the $d \times k$ matrix with columns $N_1^t, \ldots, N_k^t$, and write $\boldsymbol{N} = (N_{ij})_{ij \in [d] \times [k]}$. Then

$$\min_{\boldsymbol{y} \,:\, |\boldsymbol{y}|^2 = 1} \left\| \sum_{i=1}^{k} y_i N_i \right\|^2 = \left(\min_{\boldsymbol{y} \,:\, |\boldsymbol{y}|^2 = 1} \left\| \sum_{i=1}^{k} y_i N_i \right\| \right)^2$$

$$= \left(\min_{\boldsymbol{y} \,:\, |\boldsymbol{y}|^2 = 1} \| \boldsymbol{X} y \| \right)^2 .$$

The final quantity is just the square of the least singular value of $\boldsymbol{X}$. Theorem 1.1 of Rudelson and Vershynin [15] states the existence of absolute constants $b, B > 0$ such that for every $\varepsilon > 0$ we have

$$\mathbb{P} \left(\min_{\boldsymbol{y} \,:\, |\boldsymbol{y}| = 1} \| \boldsymbol{X} y \| \le \varepsilon (\sqrt{d} - \sqrt{k-1}) \right) \le (B\varepsilon)^{(d-k+1)} + e^{-bd} .$$

If $d \ge 4(k-1)$ then $\sqrt{d} - \sqrt{k-1} \ge \sqrt{d}/2$ and $d - k + 1 > d$. Combining the preceding probability bound with the lower bound on D, if $\varepsilon \le e^{-b}/B$ we then obtain

$$\mathbb{P} \left(D < \varepsilon^2 \frac{d}{4k} \right) < 2e^{-bd}.$$

Taking $c = (e^{-b}/B)^2$ completes the proof. $\qquad\square$

One may now easily use Proposition 4.2 to deduce part (ii) of Theorem 4.1:

Proposition 4.4. *There are absolute constants $b, C > 0$ such that the following holds. For all $p \le 1/2$, if $d \ge Cn\sqrt{\log(1/p)}$ then with probability at least $1 - e^{-bd}$ there exists $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}, s, \Phi^{-1}(1-p))$ is connected.*

Proof. Fix any tree T with vertices $[n]$, and write E for the edge set of T. By Proposition 4.2, if $d \ge (4/c) \cdot k \cdot \Phi^{-1}(1-p)$ then with probability at least $1 - e^{-bd}$ there is s such that $\Gamma(\boldsymbol{X}, s, \Phi^{-1}(1-p))$ contains T, so in particular is connected. Now simply observe that for $p \le 1/2$ we have $\Phi^{-1}(1-p) \le \sqrt{2\log(1/p)}$. $\qquad\square$

Observe that the exponentially small failure probability stipulated in Proposition 4.4 allows us to conclude that if d is at least a sufficiently large constant multiple of $n(\log n \vee \sqrt{\log(1/p)})$, then, with high probability, for *any* spanning tree of the complete graph K_n there exists $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}, s, \Phi^{-1}(1-p))$ contains that spanning tree.

Proof of Theorem 4.1, part (iii).

Let $c > 1$, $\epsilon \in (0, 1)$, and assume that $d \le (1 - \epsilon)(c - 1) \log n / \log \log n$. Let E be the event that $\Gamma(\boldsymbol{X}_n, t, s)$ is disconnected for some $s \in S^{d-1}$. Let $\mathcal{C}_\eta$ be a minimal η-cover of S^{d-1} for $\eta \in (0, 1]$ to be specified below. Then

$$E \subseteq \bigcup_{s \in \mathcal{C}_\eta} E_s \, ,$$

where E_s is the event that the graph $\bigcap_{s' : \|s - s'\| \le \eta} \Gamma(\boldsymbol{X}_n, t, s')$ is disconnected. Let $c' = c - (c-1)\epsilon/2$. Note that $1 < c' < c$. It follows from the second half of Lemma 5.4 that there exists a constant $\kappa > 0$ such that if $\eta = \kappa(1 - c'/c)/(t\sqrt{d})$, then

$$\mathbb{P}\{E_s\} \le \mathbb{P}\{G(n, c' \log n/n) \text{ is disconnected}\} \le n^{1-c'}(1 + o(1)) \, ,$$

where the second inequality follows from standard estimates for the probability that a random graph is disconnected, see Palmer [14, Section 4.3]. Bounding $|\mathcal{C}_\eta|$ by Lemma 5.1, and using the fact that $t = \sqrt{2 \log n}(1 + o(1))$, we obtain that

$$
\begin{aligned}
\mathbb{P}\{E\} &\leq |\mathcal{C}_\eta| n^{1-c'}(1 + o(1)) \\
&\leq \exp\left(\frac{d \log \log n}{2} + \frac{d \log d}{2} + O(d) + (1 - c') \log n\right) \to 0 ,
\end{aligned}
$$

as desired.

Proof of Theorem 4.1, part (iv).

Recall that $p = c \log n / n$ for $c > 1$ fixed, and that $t = \Phi^{-1}(p)$. Let $0 < \epsilon < 1$, and assume that $d \geq (2 + \epsilon)(c - 1) \log n / \log \log n$. Define $\theta \in (0, \pi/2)$ by $\theta = (\log n)^{-1/(2+\varepsilon)}$, so that $\log(1/\theta) = \log \log n / (2 + \epsilon)$. Let $\mathcal{P}$ be a maximal θ-packing of S^{d-1}, that is, $\mathcal{P} \subset S^{d-1}$ is a set of maximal cardinality such that for all distinct $s, s' \in \mathcal{P}$ we have $\langle s, s' \rangle \leq \cos \theta$. By Lemma 5.2 we have that

$$
|\mathcal{P}| \geq \frac{d}{16} \theta^{-(d-1)} .
$$

It suffices to prove that for some $s \in \mathcal{P}$, $\Gamma(\boldsymbol{X}_n, s, t)$ contains an isolated vertex.

For each $s \in \mathcal{P}$, we write the number of isolated vertices in $\Gamma(\boldsymbol{X}_n, s, t)$ as

$$
N(s) = \sum_{i=1}^{n} \prod_{j:j \neq i} Z_{i,j}(s),
$$

where $Z_{i,j}(s)$ equals 1 if $\{i, j\}$ is not an edge in $\Gamma(\boldsymbol{X}_n, s, t)$ and is 0 otherwise. We use the second moment method to prove that $N \overset{\text{def}}{=} \sum_{s \in \mathcal{P}} N(s) > 0$ with high probability. This will establish the assertion of part (iv) since if $N > 0$ then there is $s \in S^{d-1}$ such that $\Gamma(\boldsymbol{X}_n, s, t)$ contains an isolated vertex.

To show that $N > 0$ with high probability, by the second moment method it suffices to prove that $\mathbb{E}N \to \infty$ and that $\mathbb{E}[N^2] = (\mathbb{E}N)^2(1 + o(1))$. First,

$$
\mathbb{E}N = |\mathcal{P}| \cdot n \cdot \mathbb{P}\{\text{vertex 1 is isolated in } G(n, p)\} = |\mathcal{P}| \cdot n(1 - p)^{n-1}.
$$

The lower bound on $|\mathcal{P}|$ and the inequality $1 - p \leq e^{-p} = n^{-c/n}$ together imply

$$
\mathbb{E}N \geq \frac{d}{16} \theta^{-(d-1)} n^{1-c},
$$

which tends to infinity by our choice of θ. We now turn to the second moment.

$$
\mathbb{E}[N^2] = \sum_{s,s' \in \mathcal{P}} \sum_{i,j \in [n]} \prod_{k:k \neq i, \ell:\ell \neq j} Z_{i,k}(s) Z_{j,\ell}(s') .
$$

When $s = s'$, separating the inner sum into diagonal and off-diagonal terms yields the identity

$$
\sum_{i,j} \prod_{k \neq i, \ell \neq j} Z_{i,k}(s) Z_{j,\ell}(s) = n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3} = n(1-p)^{n-1} \cdot [1 + (n-1)(1-p)^{n-2}].
$$

15

Let $q = \sup_{s \neq s', s, s' \in \mathcal{P}} \mathbb{P}\{Z_{i,j}(s)Z_{i,j}(s') = 1\}$ be the greatest probability that an edge is absent in both $\Gamma(\boldsymbol{X}_n, s, t)$ and $\Gamma(\boldsymbol{X}_n, s', t)$. Then when $s \neq s'$, the inner sum is bounded by

$$nq^{n-1} + n(n-1) \cdot q \cdot (1-p)^{2n-4}.$$

Combining these bounds, we obtain that

$$\mathbb{E}[N^2] \leq |\mathcal{P}|n(1-p)^{n-1} \cdot [1 + (n-1)(1-p)^{n-2}] + |\mathcal{P}|(|\mathcal{P}| - 1) \cdot [nq^{n-1} + n(n-1) \cdot q \cdot (1-p)^{2n-4}].$$

The first term on the right is at most $\mathbb{E}N(1 + \mathbb{E}N/[(1-p)|\mathcal{P}|])$. The second is at most

$$|\mathcal{P}|^2 n^2 (1-p)^{2(n-1)} \cdot \left(\frac{1}{n}\left(\frac{q}{(1-p)^2} \right)^{n-1} + \frac{q}{(1-p)^2} \right) = (\mathbb{E}N)^2 \cdot \left(\frac{1}{n}\left(\frac{q}{(1-p)^2} \right)^{n-1} + \frac{q}{(1-p)^2} \right).$$

We will show below that $q \leq (1-p)^2 \cdot (1 + o(p))$. Assuming this, the upper bounds on the two terms on the right together give

$$\frac{\mathbb{E}[N^2]}{[\mathbb{E}N]^2} \leq \frac{1}{\mathbb{E}N}\left(1 + \frac{\mathbb{E}N}{(1-p)|\mathcal{P}|} \right) + n^{o(1)-1} + \frac{(1-\epsilon)\log n}{n} \to 1,$$

as required.

To prove the bound on q, fix $s, s' \in \mathcal{P}$ such that $q = \mathbb{P}\{Z_{i,j}(s)Z_{i,j}(s') = 1\}$. Using the definition of $Z_{i,j}(s)$ and $Z_{i,j}(s')$, we have

$$q = \mathbb{P}\left\{ \{i, j\} \notin \Gamma(\boldsymbol{X}_n, s, t), \{i, j\} \notin \Gamma(\boldsymbol{X}_n, s', t), \right\}.$$

We may apply Lemma 5.5 to this quantity, noting that in our case $\theta = (\ln n)^{1/(2+\varepsilon)}$, $t = O(\sqrt{\ln n})$ and $\ln(1/tp) = (1 + o(1)) \ln n \gg \theta^{-2}$. This means that the Remark after the statement of the Lemma applies, and this gives precisely that $q \leq (1-p)^2 (1 + o(p))$, as desired. $\qquad \square$

5   Appendix

Here we gather some of the technical tools used in the paper. In the first section we summarize results involving covering and packing results of the unit sphere that are essential in dealing with the random graph process $\mathbb{G}_{d,1/2}(\boldsymbol{X}_n)$. In Section 5.2 we describe analogous results needed for studying $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$ for small values of p. These lemmas play an important role in the proof of Theorem 4.1. Finally, in Section 5.3 we collect some results on $G(n, p)$ random graphs needed in our proofs.

5.1   Covering and packing

Let $B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}\mathrm{d}t$ be the beta function, and let $I_x(a, b)$ be the incomplete beta function,

$$I_x(a, b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1}\mathrm{d}t}{B(a, b)}.$$

For $\alpha \in [0, \pi]$ and $s \in \mathbb{S}^{d-1}$, let

$$C_\alpha(s) = \{s' \in S^{d-1} : \langle s, s' \rangle \geq \cos \alpha\}$$

be the cap in S^{d-1} consisting of points at angle at most α from s. For $\alpha \leq \pi/2$ the area of this cap (see, e.g., [10]) is

$$|C_\alpha(s)| = \frac{|S^{d-1}|}{2} \cdot I_{\sin^2 \theta}\left(\frac{d-1}{2}, \frac{1}{2}\right). \tag{1}$$

We use the following standard estimate of the covering numbers of the Euclidean sphere (see, e.g., [11, Lemma 13.1.1]).

Lemma 5.1. *For any $\eta \in (0,1]$ there exists a subset $\mathcal{C}_\eta$ of S^{d-1} of size at most $(4/\eta)^d$ such that for all $s \in S^{d-1}$ there exists $s' \in \mathcal{C}_\eta$ with $\|s - s'\| \leq \eta$.*

We now provide a rough lower bound on the number of points that can be packed in S^{d-1} while keeping all pairwise angles large.

Lemma 5.2. *For any $\theta \in (0, \pi/2)$ there exists a subset $\mathcal{P}_\theta$ of S^{d-1} of size at least*

$$\frac{d}{16}\theta^{-(d-1)}$$

such that for all distinct $s, s' \in \mathcal{P}_\theta$ we have $\langle s, s' \rangle \leq \cos \theta$.

Proof. First note that it suffices to consider $\theta < 1/2$ because otherwise the first bound dominates. Consider N independent standard normal vectors $X_1, \ldots, X_N$. Then $U_i = X_i/\|X_i\|$ ($i = 1, \ldots, N$) are independent, uniformly distributed on S^{d-1}. Let

$$Z = \sum_{i=1}^{N} \mathbb{1}_{\{\min_{j:j\neq i} |\langle U_i, U_j \rangle| \leq \cos(\theta)\}}.$$

Denoting $\mathbb{P}\{| \langle U_i, U_j \rangle | > \cos(\theta)\} = \phi$,

$$\mathbb{E}Z = N(1-\phi)^N \geq N(1 - \phi N) \geq N/2$$

whenever $\phi N \leq 1/2$. Since $Z \leq N$, this implies that

$$\mathbb{P}\left\{Z \geq \frac{N}{4}\right\} \geq \frac{\mathbb{E}Z - N/4}{N - N/4} \geq \frac{1}{3}$$

and therefore there exists a packing set A of cardinality $|A| \geq N/4$ as long as $\phi N \leq 1/2$. To study ϕ, note that

$$\phi = \mathbb{P}\left\{\frac{\sum_{j=1}^{d} Y_j Y_j'}{\|Y\| \cdot \|Y'\|} > \cos(\theta)\right\}$$

where $Y = (Y_1, \ldots, Y_d), Y' = (Y_1', \ldots, Y_d')$ are independent standard normal vectors. By rotational invariance, we may replace Y' by $(\|Y'\|, 0, \ldots, 0)$, and therefore

$$
\begin{aligned}
\phi &= \mathbb{P}\left\{\frac{Y_1^2}{\|Y\|} > \cos^2(\theta)\right\} \\
&= \mathbb{P}\left\{B \leq \cos^2(\theta)\right\} \\
&\quad \text{(where } B \text{ is a Beta}(1/2, (d-1)/2) \text{ random variable)} \\
&\geq \frac{2\theta^{d-1}}{d-1}.
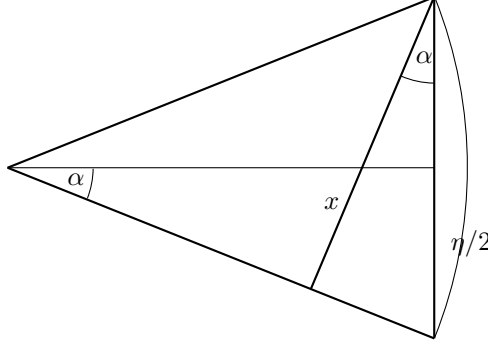\end{aligned}
$$

The result follows. $\qquad\qquad\square$

Figure 1: Since $\cos(\alpha) = \eta/2 = \sqrt{\eta^2 - x^2}/\eta$, the height of the spherical cap that only includes points at distance at least η from the equator is $1 - x = 1 - \eta\sqrt{1 - \eta^2/2}$.

The next lemma is used repeatedly in the proof of Theorem 2.1 and 3.1.

Lemma 5.3. *Fix $s' \in S^{d-1}$ and $\eta \in (0,1]$ and assume that $d \geq 12$. The probability that there exists $s \in S^{d-1}$ with $\|s - s'\| \leq \eta$ such that vertex 1 and vertex 2 are connected in $\Gamma(\boldsymbol{X}_n, s)$ but not in $\Gamma(\boldsymbol{X}_n, s')$ is at most*

$$\eta\sqrt{\frac{d}{2\pi}} \ .$$

Proof. Without loss of generality, assume that $s' = (1, 0, \ldots, 0)$. Observe that the event that there exists $s' \in S^{d-1}$ with $\|s - s'\| \leq \eta$ such that vertex 1 and vertex 2 are connected in $\Gamma(\boldsymbol{X}_n, s)$ but not in $\Gamma(\boldsymbol{X}_n, s')$ is equivalent to $X_{1,2}/\|X_{1,2}\|$ having its first component between $-\eta\sqrt{1 - \eta^2/2}$ and 0 (see Figure 1). Letting $Z = (Z_1, \ldots, Z_d)$ be a standard normal vector in $\mathbb{R}^d$, the probability of this is

$$
\begin{aligned}
\mathbb{P}\left\{\frac{Z_1}{\|Z\|} \in \left(-\eta\sqrt{1 - \eta^2/2}, 0\right)\right\} \ &\leq \ \mathbb{P}\left\{\frac{Z_1}{\|Z\|} \in (-\eta, 0)\right\} \\
&= \ \frac{1}{2}\mathbb{P}\left\{B \leq \eta^2\right\} \\
&\qquad \text{(where } B \text{ is a Beta}(1/2, (d-1)/2) \text{ random variable)} \\
&= \ \frac{1}{2}I_{\eta^2}(1/2, (d-1)/2) \\
&\leq \ \frac{1}{2B(1/2, (d-1)/2)}\int_0^{\eta^2} x^{-1/2}dx \\
&= \ \frac{\eta}{2B(1/2, (d-1)/2)} \\
&\leq \ \eta\sqrt{\frac{d-1}{2\pi}} \ . \quad \square
\end{aligned}
$$

18

5.2 Auxiliary results for $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$

In this section we develop some of the main tools for dealing with the random graph process $\mathbb{G}_{d,p}(\boldsymbol{X}_n)$. We assume throughout the section that

$$p := 1 - \Phi(t) \leq \frac{1}{2}. \tag{2}$$

Recall from the start of Section 5.1 that $C_\alpha(s)$ denotes the spherical cap consisting of all unit vectors with an angle of $\leq \alpha$ with s. We will use the following expressions for $C_\alpha(s)$:

$$\begin{aligned} C_\alpha(s) &= \{s' \in S^{d-1} : \|s - s'\|^2 \leq 2\,(1 - \cos\alpha)\} \\ &= \{s\cos\theta + w\sin\theta : w \in S^{d-1} \cap \{v\}^\perp, 0 \leq \theta \leq \alpha\}. \end{aligned} \tag{3}$$

We are interested in studying the graphs $\Gamma(\boldsymbol{X}_n, s', t)$, for all $s' \in C_\alpha(s)$ simultaneously.

Lemma 5.4. *There exists a constant $c > 0$ such that, for all $\varepsilon \in (0, 1/2)$, if $t \geq 0$ and p are as in (2),*

$$0 \leq \alpha \leq \frac{\pi}{2}, \ \tan\alpha \leq \frac{\varepsilon}{(t \vee 1)\sqrt{d-1}},$$

then, for some universal $c > 0$, if we define $\varepsilon' := \varepsilon + c\,(\varepsilon^2 + \varepsilon/(t^2 \vee 1))$,

1. *the union $\Gamma_+ := \bigcup_{s' \in C_\alpha(s)} \Gamma(\boldsymbol{X}_n, s', t)$ is stochastically dominated by $G(n, (1 + \varepsilon')\,p)$;*

2. *the intersection $\Gamma_- := \bigcap_{s' \in C_\alpha(s)} \Gamma(\boldsymbol{X}_n, s', t)$ stochastically dominates by $G(n, (1 - \varepsilon')\,p)$.*

Proof. The *first step* in this argument is to note that the edges of both Γ_+ and Γ_- are independent. To see this, just notice that, for any $\{i, j\} \in \binom{[n]}{2}$, the event that $\{i, j\}$ is an edge in $\Gamma_\pm$ depends on $\boldsymbol{X}_n$ only through $X_{i,j}$. More specifically,

$$\begin{aligned} \{i, j\} \in \Gamma_+ &\Leftrightarrow \exists s' \in C_\alpha(s) : \langle X_{i,j}, s' \rangle \geq t; \\ \{i, j\} \in \Gamma_- &\Leftrightarrow \forall s' \in C_\alpha(s) : \langle X_{i,j}, s \rangle \geq t. \end{aligned}$$

The main consequence of independence is that we will be done once we show that

$$(1 - \varepsilon')\,p \leq \mathbb{P}\{\{i, j\} \in \Gamma_-\} \leq \mathbb{P}\{\{i, j\} \in \Gamma_+\} \leq (1 + \varepsilon')\,p. \tag{4}$$

As a *second step* in our proof, we analyze the inner product of $X_{i,j}$ with $s' = s\cos\theta + w\sin\theta \in C_\alpha(s)$ (with the same notation as in (3)). Note that

$$\langle s', X_{i,j} \rangle = N\cos\theta + \langle w, X_{i,j}^\perp \rangle \sin\theta = \cos\theta \left(N + \langle w, X_{i,j}^\perp \rangle \tan\theta\right),$$

where $N := \langle X_{i,j}, s \rangle$ and $X_{i,j}^\perp$ is the component of $X_{i,j}$ that is orthogonal to s. Crucially, the fact that $X_{i,j}$ is a standard Gaussian random vector implies that N is a standard normal random variable and $X_{i,j}^\perp$ is an independent standard normal random vector in $s^\perp$. Moreover,

$$\forall w \in S^{d-1} | \langle w, X_{i,j}^\perp \rangle | \leq \chi := \|X_{i,j}^\perp\|.$$

Since "$\theta \mapsto \tan\theta$" is increasing in $[0, \alpha]$, we conclude

$$\forall s' \in C_\alpha(s) : \langle s', X_{i,j} \rangle = \cos\theta\left(N + \Delta(s')\right), \ \text{where } |\Delta(s')| \leq (\tan\alpha)\,\chi. \tag{5}$$

19

Our *third step* is to relate the above to the events $\{\{i,j\} \in \Gamma_{\pm}\}$. On the one hand,

$$
\begin{aligned}
\{i,j\} \in \Gamma_{+} \quad &\Leftrightarrow \quad \max_{s' \in C_{\alpha}(s)} \langle s', X_{i,j} \rangle \geq t \\
&\Rightarrow \quad N + \max_{s' \in \mathcal{C}_{\alpha}(s)} \Delta(s') \geq t \quad \text{(use (5) and } 0 \leq \cos\theta \leq 1) \\
&\Rightarrow \quad N \geq t - (\tan\alpha)\,\chi,
\end{aligned}
$$

and we conclude (using the independence of N and χ) that

$$
\mathbb{P}\{\{i,j\} \in \Gamma_{+}\} \leq 1 - \mathbb{E}[\Phi(t - (\tan\alpha)\,\chi)]. \tag{6}
$$

Similarly,

$$
\begin{aligned}
\{i,j\} \in \Gamma_{-} \quad &\Leftrightarrow \quad \min_{s' \in C_{\alpha}(s)} \langle s', X_{i,j} \rangle \geq t \\
&\Leftrightarrow \quad N + \min_{s' \in \mathcal{C}_{\alpha}(s)} \Delta(s') \geq \frac{t}{\cos\alpha} \quad \text{(by (5) and } \cos\theta \geq \cos\alpha > 0) \\
&\Leftarrow \quad N \geq \frac{t}{\cos\alpha} + (\tan\alpha)\,\chi,
\end{aligned}
$$

and we conclude

$$
\mathbb{P}\{\{i,j\} \in \Gamma_{-}\} \geq \mathbb{E}\left[1 - \Phi\left(\frac{t}{\cos\alpha} + (\tan\alpha)\,\chi\right)\right]. \tag{7}
$$

The remainder of the proof splits into two cases, depending on whether or not

$$
e^{\frac{5t^2}{8}}\,(1 - \Phi(t)) \geq 1 \tag{8}
$$

Note that this condition holds if and only if $t \geq C$ for some $C > 0$, as $1 - \Phi(t) = e^{-(1+o(1))t^2/2}$ when $t \to +\infty$ and $e^{\frac{5t^2}{8}}\,(1 - \Phi(t)) = 1/2 < 1$ when $t = 0$.

Last step when (8) is violated. In this case t is bounded above, so $p > c_0$ for some positive constant $c_0 > 0$. We combine (6) and (7) with the fact that $\Phi(t)$ is $(2\pi)^{-1/2}$-Lipschitz. The upshot is that

$$
|1 - \Phi(t) - \mathbb{P}\{\{i,j\} \in \Gamma_{\pm}\}| \leq \frac{1}{\sqrt{\pi}}\left|1 - \frac{1}{\cos\alpha}\right| t + \mathbb{E}[\chi]\tan\alpha.
$$

Now χ is the norm of a $d-1$ dimensional standard normal random vector, so $\mathbb{E}[\chi] \leq \sqrt{\mathbb{E}[\chi^2]} = \sqrt{d-1}$. The choice of α implies:

$$
\left|1 - \frac{1}{\cos\alpha}\right| = O(\sin\alpha) = O\left(\frac{\varepsilon^2}{d-1}\right), \quad \text{and} \quad \tan\alpha \leq \frac{\varepsilon}{\sqrt{d-1}}.
$$

So

$$
|1 - \Phi(t) - \mathbb{P}\{\{i,j\} \in \Gamma_{\pm}\}| \leq \frac{1}{\sqrt{2\pi}}\,(c\,\varepsilon^2 + \varepsilon) \leq \left[\varepsilon + c\left(\varepsilon^2 + \frac{\varepsilon}{t^2}\right)\right] p
$$

for some universal $c > 0$.

Last step when (8) is satisfied. We start with (7) and note that we can apply Lemma 5.6 with $r := t$ and

$$h := \left(\frac{1}{\cos \alpha} - 1\right) t + (\tan \alpha) \chi \le O((\tan \alpha)^2) t + (\tan \alpha) \chi.$$

After simple calculations, this gives

$$\frac{\mathbb{P}\{\{i, j\} \in \Gamma_-\}}{1 - \Phi(t)} \ge \mathbb{E}\left[\exp(-X)\right],$$

where

$$X := O((\tan \alpha)^2)(t^2 + 1) - (t + t^{-1})(\tan \alpha)\chi - (\tan \alpha)^2 \xi^2 - O((\tan \alpha)^2)t^2.$$

By Jensen's inequality, $\mathbb{E}[e^{-X}] \ge e^{-\mathbb{E}[X]}$. Since $\mathbb{E}[\chi]^2 \le \mathbb{E}[\chi^2] = d - 1$ and $\tan \alpha = \varepsilon/t\sqrt{d-1}$ in this case,

$$\mathbb{E}[X] \le O\left(\frac{\varepsilon^2}{d-1}\right) + (1 + O(\varepsilon + t^{-2}))\,\varepsilon.$$

In other words, if we choose $c > 0$ in the statement of the theorem to be large enough, we can ensure that

$$\frac{\mathbb{P}\{\{i, j\} \in \Gamma_-\}}{1 - \Phi(t)} \ge (1 - \varepsilon').$$

We now turn to (6). Applying Lemma 5.6 below with $r := t - \chi \tan \alpha$ when $r \ge t/2$, we get

$$1 - \Phi(t - (\tan \alpha)\chi) \le e^{\left(t + \frac{2}{t}\right)(\tan \alpha)\chi + \frac{(\tan \alpha)^2 \chi^2}{2}}(1 - \Phi(t)). \tag{9}$$

In fact, the same inequality holds when $r < t/2$, i.e., $(\tan \alpha)\chi > t/2$, for in that case the right-hand side is $\ge e^{\frac{5t^2}{8}}(1 - \Phi(t)) \ge 1$ (recall that we are under the assumption (8)). So (9) always holds, and integration over χ gives

$$\frac{\mathbb{P}\{\{i, j\} \in \Gamma_+\}}{1 - \Phi(t)} \le \mathbb{E}[e^{\left(t + \frac{2}{t}\right)(\tan \alpha)\chi + \frac{(\tan \alpha)^2 \chi^2}{2}}]. \tag{10}$$

It remains to estimate the moment generating function on the right-hand side. The first step is to note that, since $\mathbb{E}[\xi]$ is the norm of a $d - 1$ dimensional standard normal vector, $\mathbb{E}[\chi] \le \mathbb{E}[\chi^2]^{1/2} = \sqrt{d-1}$. So by Cauchy Schwartz,

$$
\begin{aligned}
e^{-\left(t + \frac{2}{t}\right)(\tan \alpha)\sqrt{d-1}} \mathbb{E}[e^{\left(t + \frac{2}{t}\right)(\tan \alpha)\chi + \frac{(\tan \alpha)^2 \chi^2}{2}}] &\le \mathbb{E}[e^{\left(t + \frac{2}{t}\right)(\tan \alpha)(\chi - \mathbb{E}[\chi]) + \frac{(\tan \alpha)^2 \chi^2}{2}}] \\
&\le \sqrt{\mathbb{E}[e^{\left(2t + \frac{4}{t}\right)(\tan \alpha)(\chi - \mathbb{E}[\chi])}]\mathbb{E}[e^{(\tan \alpha)^2 \chi^2}]}. \tag{11}
\end{aligned}
$$

Next we estimate each of the two expectations on the right-hand side of the last line. In the first case we have the moment generating function of $\chi - \mathbb{E}[\chi]$, where χ is a 1-Lipschitz function of a standard Gaussian vector. A standard Gaussian concentration argument and our definition of α give

$$\mathbb{E}[e^{\left(2t + \frac{4}{t}\right)(\tan \alpha)(\chi - \mathbb{E}[\chi])}] \le e^{\frac{\left(2t + \frac{4}{t}\right)^2 (\tan \alpha)^2}{2}} \le 1 + c_0 \varepsilon^2$$

for some universal constant $c_0 > 0$. The second tem in (11) is the moment generating function of χ^2, a chi-squared random variable with $d-1$ degrees of freedom. Since $(\tan \alpha)^2 \leq \varepsilon^2/(d-1) \leq 1/2$ under our assumptions, one can compute explicitly

$$\mathbb{E}[e^{(\tan \alpha)^2 \chi^2}] = \left(\frac{1}{1 - 2(\tan \alpha)^2} \right)^{d/2} \leq 1 + c_0 \varepsilon^2$$

for a (potentially larger, but still universal $c_0 > 0$). Plugging the two estimates back into (11), we obtain

$$\mathbb{E}[e^{\left(t+\frac{2}{t}\right)(\tan \alpha)\chi + \frac{(\tan \alpha)^2 \chi^2}{2}}] \leq e^{\left(t+\frac{2}{t}\right)(\tan \alpha)\sqrt{d-1}} (1 + c_0 \varepsilon^2),$$

and the fact that $t(\tan \alpha)\sqrt{d-1} = \varepsilon$ implies that the right-hand side is $\leq 1 + \varepsilon + c(t^{-2}\varepsilon + \varepsilon^2)$ for some universal $c > 0$. Going back to (10) we see that this finishes our upper bound for $\mathbb{P}\{\{i,j\} \in \Gamma_+\}$. □

Correlations between edges and non-edges

In this case we consider $s, s' \in S^{d-1}$ and look at correlations of "edge events."

Lemma 5.5. *For any $t \geq 1$, $0 < \theta < \pi$, define*

$$\xi := 1 - \cos \theta, \; \gamma := \frac{(1 - \cos \theta)^2}{\sin \theta}.$$

Then there exists a universal constant $C > 0$ such that for $s, s' \in S^{d-1}$ such that $\langle s, s' \rangle \leq \cos \theta$, we have

$$\mathbb{P}\{\langle X_{ij}, s \rangle \geq t, \langle X_{ij}, s' \rangle \geq t\} \leq p\left[(C\,p\,t)^{2\xi+\xi^2} + e^{\gamma\,(1-\gamma)\,t^2 + \frac{\gamma}{1-\gamma} + \frac{\gamma^2 t^2}{2}}\,p\right]. \tag{12}$$

$$\mathbb{P}\{\langle X_{ij}, s \rangle < t, \langle X_{ij}, s' \rangle < t\} \leq 1 - 2p + p\left[(C\,p\,t)^{2\xi+\xi^2} + e^{\gamma\,(1-\gamma)\,t^2 + \frac{\gamma}{1-\gamma} + \frac{\gamma^2 t^2}{2}}\,p\right]$$

Remark. (NEARLY EQUAL VECTORS.) Suppose $p = o(1)$ and $\theta = o(1)$. One may check that $\gamma = (1+o(1))\,\theta^3/4$ and $\xi = (1+o(1))\,\theta^2/2$. This means that if $\theta^3\,t^2 = o(\ln(1/p))$ and $\theta^2\,\ln(1/t\,p) = \omega(1)$, then

$$\mathbb{P}\{\langle X_{ij}, s \rangle < t, \langle X_{ij}, s' \rangle < t\} \leq 1 - 2p + o(p) = (1-p)^2\,(1 + o(p)).$$

This is used in the proof of Theorem 4.1, part (iv) above.

Proof. We focus on the inequalities in (12), from which the other inqualities follow. For convenience, we write $\eta := \cos \theta$ and note that

$$\eta = 1 - \xi, \; \text{so } \gamma = 1 - \frac{1 - (1+\xi)\eta}{\sqrt{1-\eta^2}}. \tag{13}$$

Moreover, $0 < \gamma < 1$: the first inequality is obvious, and the second follows from the fact that

$$0 < \theta < \frac{\pi}{2} \Rightarrow 0 < \gamma = \frac{(1 - \cos \theta)^2}{\sin \theta} < \frac{(1 - \cos \theta)\,(1 + \cos \theta)}{\sin \theta} = \frac{1 - \cos^2 \theta}{\sin \theta} = \sin \theta < 1.$$

Let E denote the event in (12). The properties of standard Gaussian vectors imply

$$\mathbb{P}\{E\} = \mathbb{P}(\{N_1 \geq t\} \cap \{\eta N_1 + \sqrt{1-\eta^2}\,N_2 \geq t\})$$

22

where N_1, N_2 are independent standard normal random variables. In particular, we can upper bound

$$\mathbb{P}\{E\} \leq \mathbb{P}\{N_1 \geq (1+\xi)\,t\} + \mathbb{P}\{N_1 \geq t\}\,\mathbb{P}\left\{N_2 \geq \left(\frac{1-(1+\xi)\eta}{\sqrt{1-\eta^2}}\right)t\right\}, \tag{14}$$

The first term in the right-hand side is $1 - \Phi(t+\xi t) \leq e^{-\frac{\xi^2 t}{2}-\xi t^2}\,(1-\Phi(t)) = e^{-\frac{2\xi+\xi^2}{2}\,t^2}\,(1-\Phi(t))$ by Lemma 5.6. The fact that

$$\lim_{t\to+\infty} \frac{(1-\Phi(t))}{e^{-t^2/2}/(t\sqrt{2\pi})} = 1,$$

implies that, for $t > 1$, the ratio $e^{-t^2/2}/p$ is bounded by a $C\,t$, $C > 0$ a constant. We conclude

$$\mathbb{P}\{N_1 \geq (1+\xi)\,t\} \leq p\,(e^{-t^2/2})^{2\xi+\xi^2} \leq p\,(C\,t\,p)^{2\xi+\xi^2}. \tag{15}$$

As for the second term in the right-hand side of (14), we apply Lemma 5.6 with

$$r := \frac{t\,(1-(1+\xi)\eta)}{\sqrt{1-\eta^2}} = (1-\gamma)\,t \text{ and } h := \gamma\,t.$$

We deduce:

$$\mathbb{P}\left\{N_2 \geq \left(\frac{1-(1+\xi)\eta}{\sqrt{1-\eta^2}}\right)t\right\} = 1 - \Phi(r) \leq e^{\gamma\,(1-\gamma)\,t^2 + \frac{\gamma}{1-\gamma} + \frac{\gamma^2\,t^2}{2}}\,(1-\Phi(t)),$$

The proof finishes by combining the estimates for the right-hand side of (14). $\qquad\square$

Lemma 5.6. *If $\varepsilon \in (0, 1/2)$, $r > 0$ and $h \geq 0$,*

$$e^{-h\,r-\frac{h}{r}-\frac{h^2}{2}} \leq \frac{1-\Phi(r+h)}{1-\Phi(r)} \leq e^{-h\,r-\frac{h^2}{2}}.$$

Proof. We first show the upper bound, namely:

$$\forall r, h > 0 : 1 - \Phi(r+h) \leq e^{-r\,h-\frac{h^2}{2}}\,(1-\Phi(r)). \tag{16}$$

To see this, we note that:

$$
\begin{aligned}
1 - \Phi(r+h) &= \int_0^{+\infty} \frac{e^{-\frac{(x+r+h)^2}{2}}}{\sqrt{2\pi}}\,dx \\
&= \int_0^{+\infty} \frac{e^{-\frac{(x+r)^2}{2}}}{\sqrt{2\pi}}\,e^{-\left(x+r+\frac{h}{2}\right)h}\,dx \\
&\leq \int_0^{+\infty} \frac{e^{-\frac{(x+r)^2}{2}}}{\sqrt{2\pi}}\,e^{-r\,h-\frac{h^2}{2}}\,dx \\
&= [1-\Phi(r)]\,e^{-r\,h-\frac{h^2}{2}}.
\end{aligned}
$$

23

To continue, we go back to the formula

$$1 - \Phi(r + h) = \left(\int_0^{+\infty} \frac{e^{-\frac{(x+r)^2}{2}} e^{-(x+r)h}}{\sqrt{2\pi}} \, dx \right) e^{-\frac{h^2}{2}},$$

which is clearly related to

$$1 - \Phi(r) = \int_0^{+\infty} \frac{e^{-\frac{(x+r)^2}{2}}}{\sqrt{2\pi}} \, dx.$$

In fact, inspection reveals that

$$\frac{1 - \Phi(r + h)}{1 - \Phi(r)} = e^{-\frac{h^2}{2}} \, \mathbb{E}[e^{-hN} \mid N \geq r].$$

Using Jensen's inequality, we have

$$\frac{1 - \Phi(r + h)}{1 - \Phi(r)} \geq e^{-\frac{h^2}{2}} e^{-h \, \mathbb{E}[N \mid N \geq r]},$$

and (16) means that $\mathbb{P}\{N - r \geq t \mid N \geq r\} \leq e^{-tr}$, so $\mathbb{E}[N \mid N \geq r] \leq r + \frac{1}{r}$. We deduce:

$$\frac{1 - \Phi(r + h)}{1 - \Phi(r)} \geq e^{-\frac{h^2}{2}} e^{-h \, r - \frac{h}{r}},$$

as desired. $\qquad \qquad \square$

5.3 Random graph lemmas

Here we collect some results on random graphs that we need in the arguments. In the proof of Theorem 2.1 we use the following lower tail estimate of the clique number of an Erdős-Rényi random graph that follows from a standard use of Janson's inequality.

Lemma 5.7. *Let N_k denote the number of cliques of size k of a $G(n, 1/2 - \alpha_n)$ Erdős-Rényi random graph where $0 \leq \alpha_n \leq 1/n$ and let $\delta > 2$. Denote $\omega = 2\log_2 n - 2\log_2 \log_2 n + 2\log_2 e - 1$. If $k = \lfloor \omega - \delta \rfloor$, then there exists a constant C' such that for all n,*

$$\mathbb{P}\{N_k = 0\} \leq \exp\left(\frac{-C' n^2}{(\log_2 n)^8} \right).$$

Proof. Write $p = 1/2 - \alpha_n$ and define $\omega_p = 2\log_{1/p} n - 2\log_{1/p}\log_{1/p} n + 2\log_{1/p}(e/2) + 1$. We use Janson's inequality ([9, Theorem 2.18]) which implies that

$$\mathbb{P}\{N_k = 0\} \leq \exp\left(\frac{-(\mathbb{E}N_k)^2}{\Delta} \right),$$

where $\mathbb{E}N_k = \binom{n}{k} p^{\binom{k}{2}}$ and

$$\Delta = \sum_{j=2}^{k} \binom{n}{k}\binom{k}{j}\binom{n-k}{k-j} p^{2\binom{k-j}{2} - \binom{j}{2} - 2j(k-j)}.$$

24

To bound the ratio $\Delta/(\mathbb{E}N_k)^2$, we may repeat the calculations of Matula's theorem on the 2-point concentration of the clique number ([12]), as in Palmer [14, Section 5.3].

Let $\beta = \log_{1/p}(3\log_{1/p} n)/\log_{1/p} n$ and define $m = \lfloor \beta k \rfloor$ Then we split the sum

$$\frac{\Delta}{(\mathbb{E}N_k)^2} = \sum_{j=m}^{k} \frac{\binom{k}{j}\binom{n-k}{k-j}}{\binom{n}{k}} p^{-\binom{j}{2}} + \sum_{j=2}^{m-1} \frac{\binom{k}{j}\binom{n-k}{k-j}}{\binom{n}{k}} p^{-\binom{j}{2}} .$$

To bound the first term, we write

$$\sum_{j=m}^{k} \frac{\binom{k}{j}\binom{n-k}{k-j}}{\binom{n}{k}} p^{-\binom{j}{2}} = \frac{F(m)}{\mathbb{E}N_k} ,$$

where $F(m) = \sum_{j=m}^{k} \binom{k}{j}\binom{n-k}{k-j} p^{-\binom{j}{2}+\binom{k}{2}}$. Now if $k = \lfloor \omega_p - \delta \rfloor$ for some $\delta \in (0, \omega_p)$, then the computations in Palmer [14, pp.77–78] show that

$$F(m) \le \sum_{j=0}^{\infty} \left(\frac{kn\sqrt{1/p}}{p^{-k(1+\beta)/2}} \right)^j ,$$

which is bounded whenever

$$\frac{kn\sqrt{(1/p)}}{p^{-k(1+\beta)/2}} = o(1) .$$

This is guaranteed by our choice of $\beta = \log_{1/p}(3\log_{1/p} n)/\log_{1/p} n$. Hence, the first term is bounded by

$$\frac{F(m)}{\mathbb{E}N_k} = O(1)\sqrt{k}p^{k\delta/2} .$$

For the second term, once again just like in [14], note that

$$\begin{aligned}
\sum_{j=2}^{m-1} \frac{\binom{k}{j}\binom{n-k}{k-j}}{\binom{n}{k}} p^{-\binom{j}{2}} &\le O(1) \sum_{j=2}^{m-1} \frac{k^{2j}}{n^j} p^{-\binom{j}{2}} \\
&\le O(1) \sum_{j=2}^{m-1} \left(\frac{kp^{-m/2}}{n} \right)^j \\
&\le O(1) \sum_{j=2}^{m-1} \left(\frac{2(\log_{1/p} n)^4}{n} \right)^j \\
&\le O\left(\frac{(\log_{1/p} n)^8}{n^2} \right) .
\end{aligned}$$

Putting everything together, we have that there exist constants C, C' such that for $k = \lfloor \omega_p - \delta \rfloor$,

$$\mathbb{P}\{N_k = 0\} \le \exp\left(-C\left(\frac{(\log_{1/p} n)^8}{n^2} + p^{k\delta/2}\sqrt{k} \right)^{-1} \right) \le \exp\left(\frac{-C'n^2}{(\log_2 n)^8} \right) ,$$

whenever $\delta > 2$. Noting that $\omega_p = \omega + O(\alpha_n \log n)$ completes the proof. $\qquad\square$

Part (iii) of Theorem 3.1 crucially hinges on the following interesting result of Alon and Sudakov [2] on the "resilience" of the chromatic number of a $G(n, 1/2)$ random graph. The form of the theorem cited here does not explicitly appear in [2] but the estimates for the probability of failure follow by a simple inspection of the proof of their Theorem 1.2.

Proposition 5.8. ([2, THEOREM 1.2]). *There exist positive constants c_1, c_2 such that the following holds. Let $\epsilon > 0$ and let G be a $G(n, 1/2)$ random graph. With probability at least $1 - \exp(c_1 n^2/(\log n)^4)$, for every collection E of at most $c_2 \epsilon^2 n^2/(\log_2 n)^2$ edges, the chromatic number of $G \cup E$ is at most $(1 + \epsilon)n/(2 \log_2 n)$.*

The final lemma is used in proving part (i) of Theorem 4.1.

Lemma 5.9. *Fix $c \in (0, 1)$. With $p = c \log n/n$, let N be the number of isolated vertices in $G(n, p)$. Then for n large, $\mathbb{P}(N = 0) \leq \exp(-n^{1-c}/3)$.*

Proof. The following approach is borrowed from O'Connell [13]. Fix $q = 1 - \sqrt{1 - p}$ and let $D(n, q)$ be the random directed graph with vertices $[n]$ in which each oriented edge ij appears independently with probability q. Write I for the number of vertices of $D(n, q)$ with no incoming edges, and M for the number of isolated vertices in $D(n, q)$, with no incoming or outgoing edges. Then M and N have the same distribution. Next, observe that I has law $\text{Bin}\left(n, (1-q)^{n-1}\right) = \text{Bin}\left(n, (1-p)^{(n-1)/2}\right)$. Furthermore, conditional on I,

$$M \stackrel{\mathrm{d}}{=} \text{Bin}\left(I, (1-p)^{(n-I)/2}\right).$$

It follows that

$$\mathbb{P}(N = 0) = \mathbb{P}(M = 0)$$

$$\leq \mathbb{P}(|I - \mathbb{E}I| > \mathbb{E}I/2) + \sup_{k \in (1/2)\mathbb{E}I, (3/2)\mathbb{E}I} \mathbb{P}(\text{Bin}\left(k, (1-p)^{(n-k)/2}\right) = 0). \qquad (17)$$

For the first term, a Chernoff bound gives

$$\mathbb{P}(|I - \mathbb{E}I| > \mathbb{E}I/2) \leq 2e^{-\mathbb{E}I/10} = 2e^{-n(1-p)^{(n-1)/2}/10} = e^{-(1+o(1))n^{1-c/2}/10}, \qquad (18)$$

where the last inequality holds since $(1 - p)^{(n-1)/2} = (1 + o(1)n^{-c/2}$. Next, fix k as in the above supremum. For such k we have $p(n - k) = c \log n + O(\log n/n^{c/2})$. Using this fact and that $1 - p \geq e^{-p-p^2}$ for p small yields

$$\mathbb{P}(\text{Bin}\left(k, (1-p)^{(n-k)/2}\right) = 0) = (1 - (1-p)^{(n-k)/2})^k$$

$$\leq \exp\left(-k(1-p)^{(n-k)/2}\right)$$

$$= \exp\left(-ke^{-(p+p^2)(n-k)/2}\right)$$

$$= \exp\left(-(1+o(1))kn^{-c/2}\right).$$

Using that $1 - p \geq e^{-p-p^2}$ a second time gives

$$k \geq \mathbb{E}I/2 = n(1-p)^{(n-1)/2}/2 \geq (1+o(1))ne^{-np/2}/2 = (1+o(1))n^{1-c/2}/2.$$

The two preceding inequalities together imply that

$$\mathbb{P}(\text{Bin}\left(k, (1-p)^{(n-k)/2}\right) = 0) \leq \exp\left(-(1/2+o(1)) \cdot n^{1-c}\right).$$

Using this bound and (18) in the inequality (17), the result follows easily. $\qquad \square$

References

[1] N. Alon and J.H. Spencer. *The Probabilistic Method.* Wiley, New York, 1992.

[2] N. Alon and B. Sudakov. Increasing the chromatic number of a random graph. *Journal of Combinatorics*, pages 345–356, 2010.

[3] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8:49–55, 1988.

[4] B. Bollobás. *Random graphs.* Cambridge University Press, Cambridge, UK, 2001.

[5] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities:A Nonasymptotic Theory of Independence.* Oxford University Press, 2013.

[6] T.M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, 14:326–334, 1965.

[7] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

[8] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[9] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs.* John Wiley, New York, 2000.

[10] S. Li. Concise formulas for the area and volume of a hyperspherical cap. *Asian J. Math. Stat.*, 4(1):66–70, 2011.

[11] J. Matoušek. *Lectures on Discrete Geometry.* Springer, 2002.

[12] D.W. Matula. Employee party problem. In *Notices of the American Mathematical Society*, volume 19, pages A382–A382, 1972.

[13] N. O'Connell. Some large deviation results for sparse random graphs. *Probab. Theory Related Fields*, 110(3):277–285, 1998.

[14] E.M. Palmer. *Graphical Evolution.* John Wiley & Sons, New York, 1985.

[15] M. Rudelson and R. Vershynin. Smallest singular value of a random rectangular matrix. *Communications in Pure and Applied Mathematics*, 62(12):1707–1739, 2009.

[16] L. Schläffli. *Gesammelte Mathematische Abhandlungen.* Birkhäuser-Verlag, Basel, 1950.

[17] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica*, 7:374–384, 1987.

[18] J.E. Steif. A survey of dynamical percolation. In *Fractal Geometry and Stochastics IV*, pages 145–174. Springer, 2009.

[19] V.N. Vapnik and A.Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16:264–280, 1971.

[20] V.N. Vapnik and A.Ya. Chervonenkis. *Theory of Pattern Recognition.* Nauka, Moscow, 1974. (in Russian); German translation: *Theorie der Zeichenerkennung*, Akademie Verlag, Berlin, 1979.

[21] V.N. Vapnik and A.Ya. Chervonenkis. Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory of Probability and its Applications*, 26:821–832, 1981.